

Encuesta Nacional de Ingresos y Gastos de los Hogares 2018

ENIGH

Nueva Serie

Diseño muestral



Instituto Nacional de Estadística y Geografía

Encuesta Nacional de Ingresos y Gastos de los Hogares 2018

ENIGH

**Nueva serie
Diseño muestral**



Obras complementarias publicadas por el INEGI sobre el tema:

Encuesta Nacional de Ingresos y Gastos de los Hogares 2016, ENIGH, Nueva serie, Diseño muestral.

Catalogación en la fuente INEGI:

640.4201 Encuesta Nacional de Ingresos y Gastos de los Hogares (2018).
Encuesta Nacional de Ingresos y Gastos de los Hogares 2018 : ENIGH : nueva serie : diseño muestral / Instituto Nacional de Estadística y Geografía.-- México : INEGI, c2019.

20 p.

1. Hogar - México - Encuestas - Metodología. 2. Economía doméstica. I. Instituto Nacional de Estadística y Geografía (México).

Conociendo México

01 800 111 4634

www.inegi.org.mx

atencion.usuarios@inegi.org.mx

 INEGI Informa  @INEGI_INFORMA

DR © 2019, **Instituto Nacional de Estadística y Geografía**

Edificio Sede

Avenida Héroe de Nacozari Sur 2301

Fraccionamiento Jardines del Parque, 20276 Aguascalientes,

Aguascalientes, Aguascalientes, entre la calle INEGI,

Avenida del Lago y Avenida Paseo de las Garzas.

Presentación

El **Instituto Nacional de Estadística y Geografía (INEGI)** presenta seis documentos metodológicos en los que se exponen las principales características de la **Encuesta Nacional de Ingresos y Gastos de los Hogares 2018 (ENIGH 2018)**.

En este documento, correspondiente al **Diseño muestral**, se describen las actividades que constituyen el proceso de obtención y tratamiento estadístico de la muestra, que inicia con la determinación de la cobertura y dominios de interés, para los cuales se pueden obtener estimaciones estadísticamente confiables, en función de los cuales se calcula el tamaño de la muestra y se determina su distribución, seguida de la selección de las unidades de muestreo.

Como parte del tratamiento, se describe el cálculo de las probabilidades de selección y su empleo para asignar el peso que debe tener cada unidad seleccionada. Finalmente se presentan las expresiones empleadas para obtener las estimaciones y sus medidas de calidad.

Índice

Diseño muestral	1
1. Objetivo de la encuesta	1
2. Población objetivo	1
3. Cobertura geográfica	1
4. Marco de la encuesta	1
4.1 Formación de las unidades primarias de muestreo (UPM)	1
4.1.1 En urbano alto	1
4.1.2 En complemento urbano	2
4.1.3 En rural	2
4.2 Estratificación	2
4.3 Selección de las UPM de la muestra maestra	2
5. Tamaño de la muestra	3
6. Afijación de la muestra	4
7. Selección de la muestra	5
7.1 En urbano alto	5
7.2 En complemento urbano	6
7.3 En rural	7
8. Ajuste a los factores de expansión	8
8.1 Ajuste por no respuesta	8
8.2 Ajuste por proyección	9
9. Estimadores	9
10. Estimación de errores de muestreo	10
11. Homologación de la Semaforización para los Umbrales de Indicadores de precisión estadística	11
Anexo	
A. Indicadores empleados en la estratificación de los conglomerados de la Muestra Maestra	15
B. Relación de las zonas dentro del ámbito urbano alto	16
C. Distribución de la muestra en viviendas por entidad según dominio de estudio para la Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH) 2018	18

Diseño muestral

1. Objetivo de la encuesta

El objetivo de la ENIGH-2018 es obtener información estadística sobre los ingresos y gastos de los hogares, así como las actividades económicas desarrolladas por los miembros de los hogares.

2. Población objetivo

La encuesta está dirigida a los hogares del territorio nacional.

3. Cobertura geográfica

La encuesta está diseñada para dar resultados a nivel nacional y entidad federativa, con cortes urbano y rural.

4. Marco de la encuesta

El diseño de la muestra para la ENIGH-2018 se caracteriza por ser probabilístico; en consecuencia, los resultados obtenidos de la encuesta se generalizan a toda la población del dominio de estudio; a su vez es bietápico, estratificado y por conglomerados, donde la unidad última de selección es la vivienda y la unidad de observación es el hogar.

La submuestra de ENIGH-2018 se seleccionó a partir de la muestra maestra 2012 del INEGI, esta muestra maestra se diseñó y seleccionó del Marco Maestro de Muestreo 2012 (MMM) el cual se conformó de conglomerados de viviendas llamados Unidades Primarias de Muestreo (UPM), contruidos a partir de la información cartográfica y demográfica que se obtuvo del Censo de Población y Vivienda 2010. La muestra maestra permite la selección de submuestras para todas las encuestas en viviendas que realiza el INEGI; su diseño es probabilístico, estratificado, unietápico y por conglomerados, pues es en ellos donde se seleccionaron, en una segunda etapa, las viviendas que integran las submuestras de las diferentes encuestas. En diseño del MMM se construyó de la siguiente manera:

4.1 Formación de las unidades primarias de muestreo (UPM)

Primeramente, se construye el conjunto de UPM que cubrirá el territorio nacional.

Las unidades primarias de muestreo están constituidas por agrupaciones de viviendas con características diferenciadas dependiendo del ámbito al que pertenecen como se especifica a continuación:

4.1.1 En urbano alto

El tamaño mínimo de una UPM es de 80 viviendas habitadas y el máximo es de 160. Pueden estar formadas por:

- Una manzana.
- La unión de dos o más manzanas contiguas de la misma AGEB¹.
- La unión de dos o más manzanas contiguas de diferentes AGEB de la misma localidad.
- La unión de dos o más manzanas contiguas de diferentes localidades, que pertenezcan al mismo tamaño de localidad.

¹ Área Geoestadística Básica.

4.1.2 En complemento urbano

El tamaño mínimo de una UPM es de 160 viviendas habitadas y el máximo es de 300. Pueden estar formadas por:

- Una manzana.
- La unión de dos o más manzanas contiguas de la misma AGEB.
- La unión de dos o más manzanas contiguas de diferentes AGEB de la misma localidad.
- La unión de dos o más manzanas contiguas de diferentes AGEB y localidades, pero del mismo municipio.

4.1.3 En rural

El tamaño mínimo de una UPM es de 160 viviendas habitadas y el máximo es 300. Pueden estar formadas por:

- Una AGEB.
- Parte de una AGEB.
- La unión de dos o más AGEB colindantes del mismo municipio.
- La unión de una AGEB con una parte de otra AGEB colindante del mismo municipio.

4.2 Estratificación

Una vez construido el conjunto de UPM, estas se agrupan considerando sus características similares, o sea se estratifican.

La división política del país y la conformación de localidades diferenciadas por su tamaño, forman de manera natural una estratificación geográfica. En cada entidad federativa se distinguen tres ámbitos, divididos a su vez en zonas, como se indica en el siguiente cuadro:

Ámbito	Zona	Descripción
Urbano alto*	01 a 09	Ciudades con 100 000 o más habitantes
Complemento urbano	25	Localidades de 50 000 a 99 999 habitantes
	35	Localidades de 15 000 a 49 999 habitantes
	45	Localidades de 5 000 a 14 999 habitantes
	55	Localidades de 2 500 a 4 999 habitantes
Rural	60	Localidades menores de 2 500 habitantes

De manera paralela, se formaron cuatro estratos sociodemográficos en los que se agruparon todas las UPM del país, esta estratificación considera las características sociodemográficas de los habitantes de las viviendas, así como, las características físicas y el equipamiento de las mismas, expresadas por medio de 34 indicadores** construidos con información del Censo de Población y Vivienda 2010, para lo cual se emplearon métodos estadísticos multivariados.

De esta forma, cada UPM fue clasificada en un único estrato geográfico y uno sociodemográfico. Como resultado, se tienen un total de 683 estratos en todo el ámbito nacional.

4.3 Selección de las UPM de la muestra maestra

Las UPM de la muestra maestra fueron seleccionadas por medio de un muestreo con probabilidad proporcional al tamaño, esto es:

$$P\{U_{ehi} \in S_m\} = \frac{k_{eh} m_{ehi}}{m_{eh}}$$

* La descripción de zonas de ámbito urbano alto se presenta en el Anexo B.

** La descripción de estos indicadores se presenta en el Anexo A.

Donde:

U_{ehi} = la i -ésima UPM, del h -ésimo estrato, de la e -ésima entidad.

S_m = la muestra maestra.

k_{eh} = número de UPM en la muestra maestra, en el h -ésimo estrato, en la e -ésima entidad.

m_{ehi} = número de viviendas en la i -ésima UPM, en el h -ésimo estrato, en la e -ésima entidad en el Censo de Población y Vivienda 2010.

m_{eh} = número de viviendas en el h -ésimo estrato, en la e -ésima entidad en el Censo de Población y Vivienda 2010.

Las UPM seleccionadas forman la muestra maestra que permite seleccionar las muestras de las encuestas en hogares.

5. Tamaño de la muestra

Como resultado del ejercicio de consulta pública, se dio revisión a los criterios para el reforzamiento de la muestra de ENIGH 2018, estableciendo los siguientes:

1. Se partió del tamaño de muestra de la ENIGH 2016, como cota mínima, para cada entidad federativa por ámbito urbano y rural.
2. Se aplicó una Tasa de no respuesta diferenciada por entidad y ámbito (urbano, rural). Partiendo de la tasa de no respuesta de la ENIGH 2016, se consideró una tasa de no respuesta mínima del 15%. En aquellos casos con tasa de no respuesta superior al 15%, se consideraron los redondeos (hacia arriba o hacia abajo) a 15, 20 y 25%.

a) Tasa de no respuesta del 20%:

Ámbito urbano: Campeche, Coahuila de Zaragoza, Colima, Chihuahua, Durango, Guerrero, Hidalgo, México, Puebla, Quintana Roo, Tabasco, Tamaulipas, Tlaxcala y Veracruz.

Ámbito rural: Chihuahua.

b) Tasa de no respuesta del 25%:

Ámbito urbano: Jalisco, Nuevo León, Querétaro, San Luis Potosí y Yucatán.

Ámbito rural: Ciudad de México.

c) El resto de las entidades y ámbitos permanece en 15 por ciento.

Además de la tasa de no respuesta, el tamaño de muestra se incrementó por ámbito (urbano, rural), considerando el tamaño de muestra mínimo requerido para estimar el promedio del ingreso corriente trimestral por hogar, el promedio del ingreso por trabajo trimestral por hogar, así como el promedio del gasto corriente trimestral por hogar, con base en la ENIGH 2016, una confianza del 90%, un margen de error relativo del 15% y una Tasa de no respuesta de acuerdo al punto 2.

3. Las entidades que requirieron incrementos de muestra fueron:

Ámbito urbano: Zacatecas.

Ámbito rural: Baja California Sur, Chihuahua, Ciudad de México, Durango, Guanajuato, Guerrero, Oaxaca, Querétaro, San Luis Potosí, Tamaulipas y Veracruz.

El ámbito rural de Oaxaca permite estimar el promedio del ingreso corriente trimestral por hogar y el promedio del gasto corriente trimestral por hogar. Sin embargo, en el caso de la variable promedio del ingreso por trabajo por hogar, se requiere una muestra superior a 3 000 viviendas, las cuales no es posible cubrir debido a la insuficiencia del marco de muestreo.

El cálculo del tamaño de muestra de la ENIGH-2018 se calculó para todas las variables de interés de la encuesta, pero para describir el procedimiento del cálculo solo se presenta para el promedio del ingreso corriente total por hogar. La expresión utilizada fue la siguiente:

$$n = \frac{z^2 s^2 DEFF}{r^2 \bar{x}^2 (1 - tnr) PHV}$$

Donde:

- n = tamaño de la muestra.
- z = valor asentado en las tablas estadísticas de la distribución normal estándar para una confianza pre-fijada.
- s^2 = estimación de la varianza poblacional de la variable de interés.
- \bar{x} = estimación del promedio de la variable de interés.
- $DEFF$ = efecto de diseño definido como el cociente de la varianza en la estimación del diseño utilizado, entre la varianza obtenida considerando un muestreo aleatorio simple para un mismo tamaño de muestra.
- r = error relativo máximo aceptable.
- tnr = tasa de no respuesta máxima esperada.
- PHV = promedio de hogares por vivienda.

El tamaño de muestra para la ENIGH 2018 se calculó a nivel Entidad con ámbito urbano y rural considerando las variables y las tasas de no respuesta mencionadas anteriormente.

A nivel Entidad con ámbito urbano, para el caso de la variable Ingreso Corriente promedio trimestral se tiene una variación entre 31 554.58 y 91 003.53, con una varianza que oscila entre 609 706 543.70 y 570 110 356 234.59, y un efecto de diseño que fluctúa entre 1.09 y 4.24.

A nivel Entidad con ámbito rural la variable Ingreso Corriente promedio trimestral varía entre 14 115.33 y 44 778.03, con una varianza que oscila entre 156 406 519.01 y 12 108 216 477.28, y un efecto de diseño que fluctúa entre 1.00 y 9.43. En el anexo C se presentan los tamaños de muestra para el ámbito Entidad urbano y rural.

Integrando los tamaños de muestra a nivel nacional se tiene un tamaño de muestra de 87 826 viviendas, el cual garantiza un error de 4.485% a nivel nacional para la variable ingreso corriente promedio trimestral.

6. Afijación de la muestra

La afijación de la muestra se realizó dentro de cada entidad federativa, de manera proporcional al número de viviendas de los estratos, para lo cual se empleó la siguiente expresión:

$$n_{eh} = \frac{N_{eh}}{N_e} n_e$$

El número de UPM a seleccionar se obtiene a partir de la siguiente expresión:

$$k_{eh} = \frac{n_{eh}}{b}$$

Donde:

n_{eh} = número de viviendas en muestra en el h-ésimo estrato, en la e-ésima entidad.

n_e = número total de viviendas en muestra en la e-ésima entidad.

N_{eh} = número total de viviendas en el h-ésimo estrato, en el e-ésima entidad.

N_e = número total de viviendas en la e-ésima entidad.

k_{eh} = número de UPM en muestra en el h-ésimo estrato, en la e-ésima entidad.

b = número promedio de viviendas a seleccionar por UPM.

En el anexo C, se presenta la distribución de la muestra en viviendas, por entidad según dominio de estudio interés.

7. Selección de la muestra

La selección de la muestra fue bietápica, donde las unidades de selección de segunda etapa fueron las viviendas y la unidad de observación el hogar. Se hizo una selección independiente para cada entidad y estrato, el procedimiento varió dependiendo del ámbito, como se especifica a continuación:

7.1 En urbano alto

En el ámbito urbano alto la selección de la muestra se realizó en forma independiente por cada entidad y estrato mediante el siguiente procedimiento:

1. De las k_{eh} UPM que integran la muestra maestra en el h-ésimo estrato, en la e-ésima entidad se eligieron k_{eh}^* UPM con igual probabilidad para la EHIGN-2018.
2. En cada UPM se seleccionaron seis viviendas con igual probabilidad para la ENIGH-2018.

La probabilidad de seleccionar la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad con probabilidad proporcional al tamaño para la muestra maestra está dado por:

$$P_{1ehi} = \frac{k_{eh} m_{ehi}}{m_{eh}}$$

La probabilidad de seleccionar la i-ésima UPM_{ehi} , del h-ésimo estrato, en la e-ésima entidad para la submuestra de la ENIGH 2018 es:

$$P_{2ehi} = \frac{k_{eh}^*}{k_{eh}}$$

La probabilidad de seleccionar a cada vivienda de la i-ésima UPM_{ehi} , del h-ésimo estrato, de la e-ésima entidad es:

$$P_{3ehi} = \frac{6}{m_{ehi}^*}$$

Por lo tanto, la probabilidad total de selección de cada una de las viviendas en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad es el producto de las probabilidades de selección:

$$P_{ehi} = P_{1ehi} * P_{2ehi} * P_{3ehi}$$

$$= \frac{k_{eh} m_{ehi}}{m_{eh}} \frac{k_{eh}^*}{k_{eh}} \frac{6}{m_{ehi}^*} = \frac{6 k_{eh}^* m_{ehi}}{m_{eh} m_{ehi}^*}$$

Y su factor de expansión² está dado por:

$$F_{ehi} = \frac{m_{eh} m_{ehi}^*}{6 k_{eh}^* m_{ehi}}$$

Donde:

- k_{eh} = número de UPM seleccionadas para la muestra maestra en el h-ésimo estrato, en la e-ésima entidad.
- k_{eh}^* = número de UPM seleccionadas para la ENIGH-2018, en el h-ésimo estrato, en la e-ésima entidad.
- m_{eh} = número de viviendas en el h-ésimo estrato, en la e-ésima entidad, según Censo de Población y Vivienda 2010.
- m_{ehi} = número de viviendas en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad, según Censo de Población y Vivienda 2010.
- m_{ehi}^* = número de viviendas en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad al momento de la actualización del listado de viviendas, previo al levantamiento de la ENIGH-2018.

7.2 En complemento urbano

- De las k_{eh} UPM que integran la muestra maestra en el h-ésimo estrato, en la e-ésima entidad, se eligieron k_{eh}^* UPM con igual probabilidad para la ENIGH-2018.
- En cada UPM seleccionada, se eligieron 24 viviendas con igual probabilidad.

La probabilidad de seleccionar la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad con probabilidad proporcional al tamaño para la muestra maestra es:

$$P_{1ehi} = \frac{k_{eh} m_{ehi}}{m_{eh}}$$

La probabilidad de selección de la i-ésima UPM_{ehi}, del h-ésimo estrato, de la e-ésima entidad para la submuestra de la ENIGH 2018 es:

$$P_{2ehi} = \frac{k_{eh}^*}{k_{eh}}$$

La probabilidad de seleccionar a cada vivienda de la i-ésima UPM_{ehi}, del h-ésimo estrato, de la e-ésima entidad es:

$$P_{3ehi} = \frac{24}{m_{ehi}^*}$$

Por lo tanto, la probabilidad total de selección de cada una de las viviendas en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad es el producto de las probabilidades de selección:

²El factor de expansión se define como el inverso de la probabilidad de selección.

$$P_{ehi} = P_{1ehi} * P_{2ehi} * P_{3ehi}$$

$$= \frac{k_{eh} m_{ehi}}{m_{eh}} \frac{k_{eh}^*}{k_{eh}} \frac{24}{m_{ehi}^*} = \frac{24 k_{eh}^* m_{ehi}}{m_{eh} m_{ehi}^*}$$

Y su factor de expansión está dado por:

$$F_{ehi} = \frac{m_{eh} m_{ehi}^*}{24 k_{eh}^* m_{ehi}}$$

Donde:

- k_{eh} = número de UPM seleccionadas para la muestra maestra, en el h-ésimo estrato, en la e-ésima entidad.
- m_{ehi} = número de viviendas en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad, según Censo de Población y Vivienda 2010.
- m_{eh} = número de viviendas en el h-ésimo estrato, en la e-ésima entidad, según Censo de Población y Vivienda 2010.
- k_{eh}^* = número de UPM seleccionadas para la ENIGH-2018, en el h-ésimo estrato, en la e-ésima entidad.
- m_{ehi}^* = número total de viviendas en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad al momento de la actualización de los listados, previo al levantamiento de la ENIGH-2018.

7.3 En rural

1. De las k_{eh} UPM que integran la muestra maestra en el h-ésimo estrato, en la e-ésima entidad, se eligieron k_{eh}^* UPM con igual probabilidad para la ENIGH-2018.
2. En cada UPM seleccionada, se eligieron 24 viviendas con igual probabilidad.

La probabilidad de seleccionar la i-ésima UPM del h-ésimo estrato, de la e-ésima entidad con probabilidad proporcional al tamaño para la muestra maestra es:

$$P_{1ehi} = \frac{k_{eh} m_{ehi}}{m_{eh}}$$

La probabilidad de selección de la i-ésima UPM_{ehi} , del h-ésimo estrato, de la e-ésima entidad para la submuestra de la ENIGH 2018 es:

$$P_{2ehi} = \frac{k_{eh}^*}{k_{eh}}$$

La probabilidad de seleccionar a cada vivienda de la i-ésima UPM_{ehi} , del h-ésimo estrato, de la e-ésima entidad es:

$$P_{3ehi} = \frac{24}{m_{ehi}^*}$$

Por lo tanto, la probabilidad total de selección de cada una de las viviendas en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad es el producto de las probabilidades de selección:

$$P_{ehi} = P_{1ehi} * P_{2ehi} * P_{3ehi}$$

$$= \frac{k_{eh} m_{ehi} k_{eh}^* 24}{m_{eh} k_{eh} m_{ehi}^*} = \frac{24 k_{eh}^* m_{ehi}}{m_{eh} m_{ehi}^*}$$

En consecuencia, su factor de expansión está dado por:

$$F_{ehi} = \frac{m_{eh} m_{ehi}^*}{24 k_{eh}^* m_{ehi}}$$

Donde:

- k_{eh} = número de UPM seleccionadas para la muestra maestra, en el h-ésimo estrato, en la e-ésima entidad.
- m_{ehi} = número de viviendas en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad, según Censo de Población y Vivienda 2010.
- m_{eh} = número total de viviendas en el h-ésimo estrato, en la e-ésima entidad, según Censo de Población y Vivienda 2010.
- k_{eh}^* = número de UPM seleccionadas para la ENIGH-2018, en el h-ésimo estrato, en la e-ésima entidad.
- m_{ehi}^* = número total de viviendas en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad al momento de la actualización de los listados, previo al levantamiento de la ENIGH-2018.

8. Ajuste a los factores de expansión

Los factores de expansión elaborados conforme al procedimiento antes descrito se ajustaron para los siguientes conceptos:

8.1 Ajuste por no respuesta

El ajuste por no respuesta atribuida al informante se realiza a nivel estrato, en cada uno de los dominios, mediante la siguiente expresión:

$$F'_{ehi} = F_{ehi} \frac{\sum_{i \in h} F_{ehi} V_{ehi}}{\sum_{i \in h} F_{ehi} V_{ehi}^*}$$

Donde:

- F'_{ehi} = factor de expansión corregido por no respuesta de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad.
- F_{ehi} = factor de expansión de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad.
- V_{ehi} = número de viviendas seleccionadas en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad.
- V_{ehi}^* = número de viviendas con respuesta en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad.

8.2 Ajuste por proyección

Los factores ajustados por no respuesta se corrigen, a fin de asegurar que en cada dominio de interés de la encuesta se obtenga la población total determinada por la proyección de población determinada por CONAPO, la cual está referida al punto medio del levantamiento, mediante la siguiente expresión:

$$F_D'' = F_D' \frac{PROY_D}{PEXP_D}$$

Donde:

F_D'' = factor de expansión corregido por proyección en el dominio D.

F_D' = factor de expansión corregido por no respuesta en el dominio D.

$PROY_D$ = población en el dominio D, según la proyección de CONAPO.

$PEXP_D$ = población total a la que expande la encuesta en el dominio D.

D = es el nivel de desagregación de la población al que se realiza el ajuste por proyección.

9. Estimadores

El estimador del total de la característica X, a nivel nacional es:

$$\hat{X} = \sum_e \sum_h \sum_i F_{ehi}^{UA} \left(\sum_s \sum_\ell X_{ehis\ell}^{UA} \right) + \sum_e \sum_h \sum_i F_{ehi}^{CU} \left(\sum_s \sum_\ell X_{ehis\ell}^{CU} \right) + \sum_e \sum_h \sum_i F_{ehi}^R \left(\sum_s \sum_\ell X_{ehis\ell}^R \right)$$

Donde:

F_{ehi}^{UA} = factor de expansión final, de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad en el ámbito urbano alto.

$X_{ehis\ell}^{UA}$ = valor observado de la característica de interés X en el ℓ -ésimo hogar, en la s-ésima vivienda, en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad en el ámbito urbano alto.

F_{ehi}^{CU} = factor de expansión final de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad, del ámbito complemento urbano.

$X_{ehis\ell}^{CU}$ = valor observado de la característica X en el ℓ -ésimo hogar, en la s-ésima vivienda, en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad del ámbito complemento urbano.

F_{ehi}^R = factor de expansión final de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad del ámbito rural.

$X_{ehis\ell}^R$ = valor observado de la característica X en el ℓ -ésimo hogar, en la s-ésima vivienda, en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad del ámbito rural.

Para la estimación de proporciones, tasas y promedios se utiliza el estimador de razón:

$$\hat{R} = \frac{\hat{X}}{\hat{Y}}$$

Donde \hat{Y} se define en forma análoga a \hat{X} .

10. Estimación de errores de muestreo

Para la evaluación de los errores de muestreo de las principales estimaciones se usa el método de “Conglomerados Últimos”³, basado en que la mayor contribución a la varianza de un estimador, en un diseño polietápico, es la que se presenta entre las unidades primarias de muestreo (UPM). El término “Conglomerados Últimos” se utiliza para denotar el total de unidades en muestra de una unidad primaria de muestreo.

Para obtener las precisiones de los estimadores de razón, se aplica el método de Conglomerados Últimos conjuntamente con el método de series de Taylor, obteniéndose la siguiente fórmula para estimar la precisión de \hat{R} .

$$\hat{V}(\hat{R}_{NAL}) = \frac{1}{\hat{Y}_{NAL}^2} \sum_{e=1}^{32} \left\{ \sum_{h=1}^{L_e} \frac{k_{eh}}{k_{eh}-1} \sum_{i=1}^{k_{eh}} \left[\left(\hat{X}_{ehi} - \frac{1}{k_{eh}} \hat{X}_{eh} \right) - \hat{R}_{NAL} \left(\hat{Y}_{ehi} - \frac{1}{k_{eh}} \hat{Y}_{eh} \right) \right]^2 \right\}$$

Donde:

\hat{X}_{ehi} = total ponderado de la variable de estudio X, en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad.

\hat{X}_{eh} = total ponderado de la variable de estudio X, en el h-ésimo estrato, en la e-ésima entidad

k_{eh} = número de UPM, en el h-ésimo estrato, en la e-ésima entidad.

L_e = número de estratos en la e-ésima entidad.

\hat{Y}_{NAL}^2 = el cuadrado del estimador del total de la característica Y.

Estas definiciones son análogas para la variable de estudio Y.

La estimación de la varianza del estimador de un total, se calcula con la siguiente expresión:

$$\hat{V}(\hat{X}_{NAL}) = \sum_{e=1}^{32} \sum_{h=1}^{L_e} \frac{k_{eh}}{k_{eh}-1} \sum_{i=1}^{k_{eh}} \left(\hat{X}_{ehi} - \frac{1}{k_{eh}} \hat{X}_{eh} \right)^2$$

Las estimaciones del error estándar (EE), efecto de diseño (DEFF) y coeficiente de variación (CV) se calculan mediante las siguientes expresiones:

$$EE = \sqrt{\hat{V}(\hat{\theta})} \quad CV = \frac{\sqrt{\hat{V}(\hat{\theta})}}{\hat{\theta}} \quad DEFF = \frac{\hat{V}(\hat{\theta})}{\hat{V}(\hat{\theta})_{mas}}$$

Donde:

$\hat{\theta}$ = estimador del parámetro poblacional θ .

$\hat{V}(\hat{\theta})_{mas}$ = estimador de la varianza bajo muestreo aleatorio simple.

$\hat{V}(\hat{\theta})$ = estimador de la varianza bajo el diseño de muestreo descrito en este documento.

³ Véase Hansen, M.H. Horwitz, W.N. y Madow, W.G, Sample Survey Methods and Theory, (1953), Vol. 1 página 242.

Finalmente, el intervalo de confianza $I_{1-\alpha}$ al $100(1-\alpha)\%$, se construye de la siguiente forma:

$$I_{1-\alpha} = \left(\hat{\theta} - z_{1-\alpha/2} \sqrt{\hat{V}(\hat{\theta})}, \hat{\theta} + z_{1-\alpha/2} \sqrt{\hat{V}(\hat{\theta})} \right)$$

Donde α es el nivel de significancia.

11. Homologación de la Semaforización para los Umbrales de Indicadores de precisión estadística

Para facilitar la interpretación de las precisiones estadísticas de la información pública en tabulados, el Comité de Aseguramiento de la Calidad, en la cuarta sesión celebrada el 1 de noviembre de 2018, aprobaron los siguientes umbrales y especificaciones para la publicación en los tabulados los CV, así como su semaforización de estos.

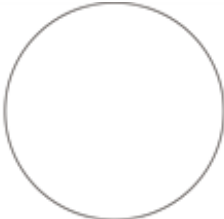
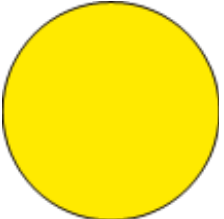
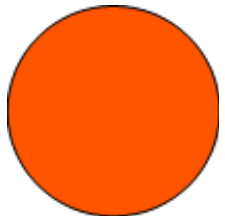
Umbrales aprobados para los CV.

Interpretación	Semaforización	Intervalo del CV
Alta	Blanco	[0%, 15%)
Moderada	Amarillo	[15%, 30%)
Baja	Naranja oscuro	$\geq 30\%$

Umbrales aprobados para el reporte de la precisión de acuerdo con el coeficiente de variación en los tabulados de resultados de los proyectos con muestreo probabilístico (acuerdo CAC-007/01/2018).

A partir del segundo trimestre de 2018, se publican los siguientes indicadores de precisión estadística en la presentación de resultados en tabulados de todas las encuestas con muestreo probabilístico del INEGI: error estándar, intervalo de confianza y coeficiente de variación (CV). Adicionalmente, se estandariza la coloración en los tabulados para indicar el nivel de precisión de las estimaciones con base en el CV. A continuación, se presenta el código RGB de los colores utilizados en la semaforización:

Parámetros RGB para la semaforización del coeficiente de variación

Alta Blanco	Moderada Amarillo	Baja Naranja
		
Rojo: 255 Verde: 255 Azul: 255	Rojo: 255 Verde: 234 Azul: 0	Rojo: 255 Verde: 84 Azul: 0

El siguiente texto explicativo aparece en cada uno de los tabulados publicados de encuestas por muestreo probabilístico.

Las estimaciones que aparecen en este cuadro están coloreadas de acuerdo con su nivel de precisión, en *Alta*, *Moderada* y *Baja*, tomando como referencia el coeficiente de variación CV (%). Una precisión *Baja* requiere un uso cauteloso de la estimación en el que se analicen las causas de la alta variabilidad y se consideren otros indicadores de precisión y confiabilidad, como el intervalo de confianza.

Nivel de precisión de las estimaciones:

Alta, CV en el rango de (0,15)

Moderada, CV en el rango de [15, 30)

Baja, CV de 30% en adelante

Anexo

A. Indicadores empleados en la estratificación de los conglomerados de la Muestra Maestra

Mnemónico	Descripción
Proporción de Poblacion	
PPSSNOSP	Que tiene derecho a recibir servicios médicos en alguna institución de salud pública o privada excepto seguro popular.
PPDER_SS	Derechohabiente a servicios de salud.
PDP3A14A	De 3 a 14 años de edad que asiste a la escuela.
PDP15A24A	De 15 a 24 años de edad que asiste a la escuela.
PDP8A14ALF	De 8 a 14 años de edad que saben leer o escribir.
PDP15YM_SE	De 15 años o más de edad que aprobaron algún grado de escolaridad diferente al nivel preescolar.
PP15PRI_CO	De 15 años o más de edad que tienen como máxima escolaridad 6 grados aprobados en primaria.
PP15SEC_CO	De 15 años o más de edad que tienen como máxima escolaridad 3 grados aprobados en secundaria.
PGDO_ESC	Grado promedio de escolaridad.
PPEA	De 12 años y más que trabajaron; tenían trabajo pero no trabajaron o; buscaron trabajo en la semana de referencia.
PPEA_F	Femenina de 12 años y más que trabajaron; tenían trabajo pero no trabajaron o; buscaron trabajo en la semana de referencia.
PTASAOCUPA	Tasa de ocupación.
TOCU12A17	No ocupada de 12 a 17 años de edad entre la población de 12 a 17 años de edad.
PPOMAYED	Ocupada de 18 y más años de edad entre la población ocupada.
Proporción de Viviendas Particulares Habitadas	
PVIVSINH	Que no tienen hacinamiento.
PVPH_PISDT	Que tienen piso de cemento o firme, madera, mosaico u otro material.
PVPH2YMASD	Que usan para dormir entre 2 y 25 cuartos.
PVPH_2MASC	Que tienen más de un cuarto.
PVPH3YMASD	Que tienen entre 3 y 25 cuartos.
PVPH_C_ELE	Que disponen de luz eléctrica.
PVPHAGUADV	Que tienen disponibilidad de agua entubada dentro de la vivienda, o fuera de la vivienda pero dentro del terreno.
PVPH_EXCSA	Que tienen excusado, retrete, sanitario, letrina u hoyo negro.
PVPHDRENAJ	Que tienen drenaje conectado a la red pública, fosa séptica, barranca, grieta, río, lago o mar.
PVDRENERED	Que disponen de drenaje conectado a la red pública.
PVEXCAGU	Que disponen de excusado con descarga directa de agua.
PVPH_CSERV	Que disponen de luz eléctrica, agua entubada dentro o fuera de la vivienda, pero dentro del terreno, así como drenaje.
PSIN_HASIN	Que no se encuentra en situación de hacinamiento a nivel manzana.
Proporción de Viviendas Particulares Habitadas que disponen de:	
PVPH_TV	Televisor.
PVPH_AUTOM	Automóvil o camioneta.
PVPH_CEL	Teléfono celular.
PVCELFIJ	Teléfono celular y teléfono fijo.
PV4ELEC	Radio, televisor, refrigerador y lavadora.
PVRADTEL	Radio y televisor.
PVPHCBIEN	Todos los bienes.

B. Relación de las zonas dentro del ámbito urbano alto

Clave	Entidad	Zona	Nombre de la zona
01	Aguascalientes	01	Aguascalientes
02	Baja California	01	Tijuana
		02	Mexicali
		03	Ensenada
03	Baja California Sur	01	La Paz
04	Campeche	01	Campeche
		02	Ciudad del Carmen
05	Coahuila de Zaragoza	01	Saltillo
		02	Torreón
		03	Monclova
		04	Piedras Negras
		05	Ciudad Acuña
06	Colima	01	Colima
		02	Manzanillo
07	Chiapas	01	Tuxtla Gutiérrez
		02	Tapachula
		03	San Cristóbal de las Casas
08	Chihuahua	01	Ciudad Juárez
		02	Chihuahua
		03	Delicias
		04	Hidalgo del Parral
		05	Cuauhtémoc
09	Ciudad de México	01	Ciudad de México
10	Durango	01	Durango
		02	Torreón
11	Guanajuato	01	León
		02	Irapuato
		03	Celaya
		04	Salamanca
		05	Guanajuato
12	Guerrero	01	Acapulco
		02	Chilpancingo
		03	Iguala
13	Hidalgo	01	Pachuca
		02	Tulancingo
14	Jalisco	01	Guadalajara
		02	Puerto Vallarta
15	México	01	Ciudad de México
		02	Toluca
16	Michoacán de Ocampo	01	Morelia
		02	Uruapan
		03	Zamora
17	Morelos	01	Cuernavaca
		02	Cuatla

B. Relación de las zonas dentro del ámbito urbano alto

Clave	Entidad	Zona	Nombre de la zona
18	Nayarit	01	Tepic
19	Nuevo León	01	Monterrey
20	Oaxaca	01	Oaxaca
		05	San Juan Bautista Tuxtepec
21	Puebla	01	Puebla
		02	Tehuacán
22	Querétaro	01	Querétaro
		02	San Juan del Río
23	Quintana Roo	01	Cancún
		02	Chetumal
		03	Playa del Carmen
24	San Luis Potosí	01	San Luis Potosí
		02	Ciudad Valles
25	Sinaloa	01	Culiacán Rosales
		02	Mazatlán
		03	Los Mochis
26	Sonora	01	Hermosillo
		02	Ciudad Obregón
		03	Heroica Nogales
		04	San Luis Río Colorado
		05	Guaymas
		06	Navojoa
27	Tabasco	01	Villahermosa
28	Tamaulipas	01	Tampico
		02	Reynosa
		03	Matamoros
		04	Nuevo Laredo
		05	Ciudad Victoria
29	Tlaxcala	01	Tlaxcala
30	Veracruz de Ignacio de la Llave	01	Veracruz
		02	Xalapa
		03	Orizaba
		04	Coatzacoalcos
		05	Poza Rica
		06	Córdoba
		07	Minatitlán
		08	Tuxpan
		09	Tampico
31	Yucatán	01	Mérida
32	Zacatecas	01	Zacatecas
		02	Fresnillo

C. Distribución de la muestra en viviendas por entidad según dominio de estudio para la Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH) 2018

Entidad		Dominio	ENIGH 2016	Propuesta Consulta	Nuevas TNR Consulta Pública	Nuevo ENIGH 2018
Clave	Nombre					
Estados Unidos Mexicanos		Nacional	80 328	87 320	0.15	87 826
		Urbano	53 076	57 720	0.15	56 052
		Rural	27 252	29 600	0.15	31 774
01	Aguascalientes	Entidad	2 616	2 620		2 640
		Urbano	1 698	1 700	0.15	1 698
		Rural	918	920	0.15	942
02	Baja California	Entidad	4 032	4 160		4 035
		Urbano	1 860	1 980	0.15	1 860
		Rural	2 172	2 180	0.15	2 175
03	Baja California Sur	Entidad	2 604	2 920		2 737
		Urbano	1 770	1 900	0.15	1 770
		Rural	834	1 020	0.15	967
04	Campeche	Entidad	2 004	2 120		2 099
		Urbano	1 524	1 620	0.20	1 620
		Rural	480	500	0.15	479
05	Coahuila de Zaragoza	Entidad	3 936	4 160		4 146
		Urbano	3 174	3 380	0.20	3 378
		Rural	762	780	0.15	768
06	Colima	Entidad	3 150	3 280		3 270
		Urbano	1 860	1 980	0.20	1 980
		Rural	1 290	1 300	0.15	1 290
07	Chiapas	Entidad	2 004	2 080		2 002
		Urbano	1 068	1 140	0.15	1 068
		Rural	936	940	0.15	934
08	Chihuahua	Entidad	3 780	4 700		4 521
		Urbano	2 634	3 000	0.20	2 802
		Rural	1 146	1 700	0.20	1 719
09	Ciudad de México	Entidad	2 004	2 800		2 485
		Urbano	1 980	2 300	0.15	1 980
		Rural	24	500	0.25	505
10	Durango	Entidad	2 316	2 420		2 646
		Urbano	1 404	1 500	0.20	1 494
		Rural	912	920	0.15	1 152
11	Guanajuato	Entidad	2 586	2 620		2 856
		Urbano	1 404	1 420	0.15	1 404
		Rural	1 182	1 200	0.15	1 452
12	Guerrero	Entidad	2 004	2 180		2 420
		Urbano	1 212	1 380	0.20	1 290
		Rural	792	800	0.15	1 130

C. Distribución de la muestra en viviendas por entidad según dominio de estudio para la Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH) 2018

Entidad		Dominio	ENIGH 2016	Propuesta Consulta	Nuevas TNR Consulta Pública	Nuevo ENIGH 2018
Clave	Nombre					
13	Hidalgo	Entidad	2 046	2 140		2 119
		Urbano	1 116	1 200	0.20	1 188
		Rural	930	940	0.15	931
14	Jalisco	Entidad	2 418	2 980		2 679
		Urbano	1 740	2 300	0.25	1 980
		Rural	678	680	0.15	699
15	Estado de México	Entidad	3 162	3 280		3 296
		Urbano	1 764	1 880	0.20	1 878
		Rural	1 398	1 400	0.15	1 418
16	Michoacán de Ocampo	Entidad	2 004	2 020		2 010
		Urbano	1 380	1 380	0.15	1 380
		Rural	624	640	0.15	630
17	Morelos	Entidad	2 442	2 560		2 467
		Urbano	1 692	1 800	0.15	1 692
		Rural	750	760	0.15	775
18	Nayarit	Entidad	2 004	2 020		2 005
		Urbano	1 404	1 420	0.15	1 404
		Rural	600	600	0.15	601
19	Nuevo León	Entidad	3 606	3 940		3 870
		Urbano	1 974	2 300	0.25	2 244
		Rural	1 632	1 640	0.15	1 626
20	Oaxaca	Entidad	2 004	2 420		2 448
		Urbano	972	980	0.15	972
		Rural	1 032	1 440	0.15	1 476
21	Puebla	Entidad	2 004	2 120		2 107
		Urbano	1 452	1 560	0.20	1 548
		Rural	552	560	0.15	559
22	Querétaro	Entidad	2 616	2 860		3 640
		Urbano	1 644	1 880	0.25	1 866
		Rural	972	980	0.15	1 774
23	Quintana Roo	Entidad	2 196	2 540		2 312
		Urbano	1 788	2 040	0.20	1 902
		Rural	408	500	0.15	410
24	San Luis Potosí	Entidad	2 004	2 200		2 568
		Urbano	1 308	1 500	0.25	1 488
		Rural	696	700	0.15	1 080
25	Sinaloa	Entidad	3 306	3 320		3 309
		Urbano	2 778	2 780	0.15	2 778
		Rural	528	540	0.15	531

C. Distribución de la muestra en viviendas por entidad según dominio de estudio para la Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH) 2018

Clave	Entidad		Dominio	ENIGH 2016	Propuesta Consulta	Nuevas TNR Consulta Pública	Nuevo ENIGH 2018
	Nombre						
26	Sonora		Entidad	2 454	2 460		2 446
			Urbano	1 740	1 740	0.15	1 740
			Rural	714	720	0.15	706
27	Tabasco		Entidad	2 004	2 180		2 084
			Urbano	1 188	1 360	0.20	1 266
			Rural	816	820	0.15	818
28	Tamaulipas		Entidad	2 178	2 380		2 377
			Urbano	1 764	1 880	0.20	1 878
			Rural	414	500	0.15	499
29	Tlaxcala		Entidad	2 004	2 380		2 137
			Urbano	1 644	1 880	0.20	1 752
			Rural	360	500	0.15	385
30	Veracruz de Ignacio de la Llave		Entidad	2 004	2 280		2 715
			Urbano	1 260	1 340	0.20	1 344
			Rural	744	940	0.15	1 371
31	Yucatán		Entidad	2 742	2 980		2 976
			Urbano	1 692	1 920	0.25	1 920
			Rural	1 050	1 060	0.15	1 056
32	Zacatecas		Entidad	2 094	2 200		2 404
			Urbano	1 188	1 280	0.15	1 488
			Rural	906	920	0.15	916